

Universality of Non-Local Boxes

Manuel Forster and Stefan Wolf

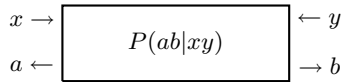
Computer Science Department, ETH Zürich, CH-8092 Zürich, Switzerland

(Dated: November 17, 2008)

One of the most fascinating consequences of quantum theory is *non-locality*, i.e., the fact that the behavior under measurements of (spatially separated) parts of a system can have a correlation unexplainable by shared classical information. Note that at the same time, these correlations are *non-signaling* and do not allow for message transmission. Popescu and Rohrlich have defined a *non-local box* as a “basic building block of non-locality” and initiated a systematic study of non-local correlations and their applications. They left open, however, whether *any* bi-partite non-signaling correlation can be simulated by such non-local boxes. We show that the answer is *yes* with respect to arbitrarily accurate approximations.

INTRODUCTION

In probability theory, the term correlation is often used to indicate a departure from independence. When two separated parts of a quantum state are measured, then the outcomes can be correlated. While this may be surprising from a physical point of view, it is much less from an information-theoretical perspective since such correlations could have arisen by randomness shared when the two particles were generated. In this note we address correlations of a stronger kind - those which are unexplainable by shared randomness. More precisely, we study correlations between the outcomes of the two ends of a bipartite input-output *system*, characterized by a conditional probability distribution $P(ab|xy)$. Let x and a stand for the input and output on the left-hand side of the system, and y and b for the corresponding values on the right-hand side.



We call such a behavior *local* if it is explainable by shared classical information and *signaling* if it allows for message transmission in either direction.

Local correlations satisfy certain linear inequalities known as *Bell inequalities* [1]; in other words, violation of such inequalities indicates non-locality. John Bell was the first to realize that entangled quantum states can have a non-local behavior under measurements if x and y are the choice of the measurements to be carried out, and a and b are the measurement results. The non-local behavior of quantum systems was highlighted by Einstein, Podolsky and Rosen [2], and what was originally (erroneously [1]) considered a witness for the incompleteness of quantum mechanics, has proven to be a very useful information theoretic resource [3].

In accordance with relativity theory, the joint behavior of quantum systems must be *non-signaling*: otherwise superluminal message transmission would be possible. A binary input-output system characterized by a conditional probability distribution $P(ab|xy)$ is *non-*

signaling if one cannot signal from one side to the other by the choice of the input.

We are interested in systems that are *neither signaling nor local*. It may not be obvious that such correlations can be defined, which are impossible to be simulated by quantum mechanics. Surprisingly, they can, and an example is the *non-local box* (NLB for short) or *Popescu-Rohrlich (PR) machine* [4]. The NLB is a system violating the Clauser-Horne-Shimony-Holt (CHSH) inequality [5] to the algebraic maximum. Its behavior is as follows: We have $x, y, a, b \in \{0, 1\}$; a and b are uniform and independent of (x, y) , but $a + b \equiv xy \pmod{2}$ always holds. Its probability distribution is, thus,

$$P_{\text{NLB}}(ab|xy) = \begin{cases} 1/2 & \text{if } xy = a + b \pmod{2}, \\ 0 & \text{otherwise.} \end{cases}$$

The PR machine cannot occur as the behavior of a quantum state [6], but it can be approximated with an accuracy of roughly 85%, whereas 75% is the local limit.

With the definition of the non-local box, Popescu and Rohrlich raised the question why quantum mechanics is not maximally non-local, that is, why quantum non-locality is not solely constrained by the non-signaling conditions. It is hence possible to study non-locality without reference to quantum mechanics, but simply as a property of joint probability distributions. The non-local box has since been considered as a possible candidate for an “atom” or “basic building block” for non-locality. Clearly, this notion makes more sense if, actually, non-local boxes allow the realization of *any* non-signaling system, that is, if two parties can non-interactively simulate any non-signaling system by local operations on shared randomness and a finite quantity of non-local boxes.

Previous Work

In [7], and independently in [8], it is proven that every non-signaling system with *binary outputs* $a, b \in \{0, 1\}$ can be simulated by local operations on a finite number of NLBs. In [9], the alternative case of *binary inputs*

$x, y \in \{0, 1\}$ is solved. In the latter case, the simulation can be made arbitrarily close, but not perfect. In certain cases perfect simulation is known to be impossible [10].

Our Contribution

We show that, on the other hand, if one is willing to accept an arbitrarily small error, then non-local boxes are universal.

Theorem 1 (Main Result). *Any bipartite non-signaling system can be approximated arbitrarily closely by shared randomness and non-local boxes.*

DEFINITIONS

A bi-partite input-output *system* is characterized by a conditional probability distribution $P(ab|xy)$ where $a \in \mathcal{A}$ and $x \in \mathcal{X}$ are the input and output one the left hand side of the system and $b \in \mathcal{B}$ and $y \in \mathcal{Y}$ the corresponding values on the right hand side, respectively. A system is *non-signaling* if it cannot be used to signal from one side to the other by the choice of the input. This means that the marginal probabilities $P(a|x)$ and $P(b|y)$ are independent of y and x , respectively, i.e.,

$$\begin{aligned} \sum_b P(ab|xy) &= \sum_b P(ab|xy') \equiv P(a|x) \quad \forall a, x, y, y', \\ \sum_a P(ab|xy) &= \sum_a P(ab|x'y) \equiv P(b|y) \quad \forall b, x, x', y. \end{aligned}$$

Using a non-signaling system a party receives its output immediately after giving its input, independently of whether the other has given its input already. Note that this is possible since the system is non-signaling. Furthermore, after a system is used once it is destroyed.

Definition 1. *A permutation f on a set S is a bijective self mapping $f : S \rightarrow S$.*

Let \mathcal{P} denote the set of all bi-partite non-signaling systems. We define another set $\mathcal{D} \subset \mathcal{P}$ as follows:

Definition 2. *Let \mathcal{D} include all non-signaling systems D_d with output set $S_d = \{0, 1, \dots, d-1\}$ and arbitrary input sets \mathcal{X}, \mathcal{Y} . For $a, b \in S_d$ and $x \in \mathcal{X}, y \in \mathcal{Y}$ its conditional probability distribution is defined as*

$$P_{D_d}(ab|xy) = \begin{cases} 1/d & \text{if } f_{xy}(a) = b, \\ 0 & \text{otherwise,} \end{cases}$$

where $f_{xy} : S \rightarrow S$ is a permutation on the output set depending on x, y . Every $D_d \in \mathcal{D}$ is fully defined by the set $\{f_{xy} : x \in \mathcal{X}, y \in \mathcal{Y}\}$.

It is not hard to see that the marginal probabilities $P_{D_d}(a|x)$ as well as $P_{D_d}(b|y)$ are uniform independently of y and x , respectively. Therefore $\mathcal{D} \subset \mathcal{P}$ holds. Refer to the NLB as a prominent example of a $D_2 \in \mathcal{D}$.

RESULTS

In this section we will prove Theorem 1. We distinguish two main parts of the proof: First (Subsection), we restrict the reasoning to the set \mathcal{D} and show that with a finite number of $D_{d-1} \in \mathcal{D}$ and shared randomness one can approximate any $D_d \in \mathcal{D}$ arbitrarily well. This induction step is anchored by the fact that any $D_2 \in \mathcal{D}$, i.e., any distributed Boolean function, can be simulated by local operations on a finite number of NLBs [11]. Second (Subsection), the universality of the set \mathcal{D} itself is proven, that is, given an arbitrary non-signaling system $P \in \mathcal{P}$ we show a protocol that simulates P by local operations on a $D_d[P] \in \mathcal{D}$ arbitrarily closely.

Induction on d

In the following we reduce the problem of simulating any given system D_d to the task of simulating a finite number of D_{d-1} . We define a simulation protocol as a number of classical operations on *resource* systems and (if needed) shared randomness. The operations are subsequently and non-interactively executed by two parties (usually called Alice and Bob) which finally simulate a certain *target* system.

Constructing D_{d-1}

Suppose that a system $D_d \in \mathcal{D}$, defined by the output set S_d and the permutations $\{f_{xy} : x \in \mathcal{X}, y \in \mathcal{Y}\}$, is given. Let therefore D_{d-1} be characterized by the output set $S_{d-1} = S_d \setminus \{d-1\}$ and input sets $\mathcal{X}' = \mathcal{X} \times S_d$, $\mathcal{Y}' = \mathcal{Y} \times S_d$. For $a, b \in S_{d-1}$ and $x' \in \mathcal{X}', y' \in \mathcal{Y}'$ its conditional probability distribution shall be

$$P_{D_{d-1}}(ab|x'y') = \begin{cases} 1/(d-1) & \text{if } g_{x'y'}(a) = b, \\ 0 & \text{otherwise.} \end{cases}$$

The function $g_{x'y'}$ is fixed by the inputs $x' = (x, a_0)$ and $y' = (y, b_0)$ as

$$g_{x'y'}(a) = \begin{cases} r_{b_0}^{-1}(f_{xy}(a_0)) & \text{if } f_{xy}(a) = b_0 \wedge a \neq a_0, \\ r_{b_0}^{-1}(f_{xy}(r_{a_0}(a))) & \text{otherwise,} \end{cases} \quad (1)$$

where $r_{a_0} : S_{d-1} \rightarrow S_d \setminus \{a_0\}$ and $r_{b_0} : S_{d-1} \rightarrow S_d \setminus \{b_0\}$ are bijections

$$r_{a_0}(a) = \begin{cases} d-1 & \text{if } a = a_0, \\ a & \text{otherwise,} \end{cases} \quad (2)$$

$$r_{b_0}(b) = \begin{cases} d-1 & \text{if } b = b_0, \\ b & \text{otherwise.} \end{cases} \quad (3)$$

The fact that D_{d-1} as defined above is a valid system in \mathcal{D} is proven by the following corollary:

Corollary 1. *It holds that $D_{d-1} \in \mathcal{D}$.*

Proof. Observe that the system D_{d-1} is fully characterized by the set $\{g_{x'y'} : x' \in \mathcal{X}', y' \in \mathcal{Y}'\}$. Therefore, we will prove that $g_{x'y'}$ is a permutation on S_{d-1} .

Given are arbitrary $x \in \mathcal{X}$, $a_0 \in S_d$ and $y \in \mathcal{Y}$, $b_0 \in S_d$. Since $r_{a_0}(a) \neq a_0$ for all $a \in S_{d-1}$ (2) we need only distinguish the following two cases in (1):

$$\begin{aligned} \forall a, a' \in S_{d-1} : g_{xa_0, yb_0}(a) &= g_{xa_0, yb_0}(a') \\ \Rightarrow r_{b_0}^{-1}(f_{xy}(a_0)) &= r_{b_0}^{-1}(f_{xy}(a_0)) \\ \Rightarrow f_{xy}(a) &= b_0 \wedge a \neq a_0 \wedge f_{xy}(a') = b_0 \wedge a' \neq a_0 \\ \Rightarrow a &= a'. \end{aligned}$$

Here we used that f_{xy} is a permutation and, therefore, $f_{xy}(a) = b_0$ for exactly one input value a . Otherwise:

$$\begin{aligned} \forall a, a' \in S_{d-1} : g_{xa_0, yb_0}(a) &= g_{xa_0, yb_0}(a') \\ \Rightarrow r_{b_0}^{-1}(f_{xy}(r_{a_0}(a))) &= r_{b_0}^{-1}(f_{xy}(r_{a_0}(a'))) \\ \Rightarrow a &= a', \end{aligned}$$

since $r_{b_0}^{-1}, r_{a_0}$ are bijections. So g_{xa_0, yb_0} is injective from a finite set to itself — a permutation on S_{d-1} . \square

The Simulation Protocol

The following protocol consists of a finite number of rounds. Each round includes four steps subsequently and non-interactively executed by Alice and Bob. These are:

1. Let $a_0, b_0 \in_{uar} S_d$ in the initial round, otherwise $a_0 = a_t, b_0 = b_t$. Alice holds a_0 and Bob b_0 .
2. Let s denote a random bit shared by Alice and Bob, such that $P(s = 0) = 1/d$.
3. System D_{d-1} is locally accessed by Alice and Bob through the functions $D_{d-1}^A : \mathcal{X} \times S_d \rightarrow S_{d-1}$ and $D_{d-1}^B : \mathcal{Y} \times S_d \rightarrow S_{d-1}$, respectively. Therefore, on input (x, a_0) Alice obtains $a = D_{d-1}^A(x, a_0)$ and Bob $b = D_{d-1}^B(y, b_0)$ on input (y, b_0) , respectively. From the definition of D_{d-1} we have that always

$$g_{xa_0, yb_0}(a) = b. \quad (4)$$

4. Let the local computation of temporary outputs $a_t, b_t \in S_d$ be defined as

$$a_t = (1 - s)a_0 + sr_{a_0}(a) \quad (5)$$

for Alice and

$$b_t = (1 - s)b_0 + sr_{b_0}(b) \quad (6)$$

for Bob.

ALICE	SIM(n)	BOB
$x \in X$	inputs	$y \in Y$
1st round		
$a_0 \in_{uar} S_d$		$b_0 \in_{uar} S_d$
$s \in \{0, 1\}$	$P(s = 0) = \frac{1}{d}$	$s \in \{0, 1\}$
$a = D_{d-1}^A(x, a_0)$	D_{d-1}	$b = D_{d-1}^B(y, b_0)$
$a_t = (1 - s)a_0$		$b_t = (1 - s)b_0$
$+sr_{a_0}(a)$		$+sr_{b_0}(b)$
2nd round		
$a_0 = a$		$b_0 = b$
$s \in \{0, 1\}$	$P(s = 0) = \frac{1}{d}$	$s \in \{0, 1\}$
$a = D_{d-1}^A(x, a_0)$	D_{d-1}	$b = D_{d-1}^B(y, b_0)$
$a_t = (1 - s)a_0$		$b_t = (1 - s)b_0$
$+sr_{a_0}(a)$		$+sr_{b_0}(b)$
\vdots	\vdots	\vdots
nth round		
$a_0 = a$		$b_0 = b$
$s \in \{0, 1\}$	$P(s = 0) = \frac{1}{d}$	$s \in \{0, 1\}$
$a = D_{d-1}^A(x, a_0)$	D_{d-1}	$b = D_{d-1}^B(y, b_0)$
$a_t = (1 - s)a_0$	outputs	$b_t = (1 - s)b_0$
$+sr_{a_0}(a)$		$+sr_{b_0}(b)$

FIG. 1: The simulation of D_d by local operations on n D_{d-1} .

The next round starts with assignments $a_0 = a_t, b_0 = b_t$ and then proceeds with the second step and so on. After the last round the parties output a_t, b_t as the final outputs of the simulation.

(The relabeling functions r_{a_0}, r_{b_0} are necessary because the set S_{d-1} , on which $g_{x'y'}$ is defined, obviously lacks the element $d - 1$, which is one of D_d 's outputs. On the other hand we have already correlated the outputs a_0, b_0 in the case $s = 0$, so these outputs can serve as replacements for $d - 1$ and $f_{xy}(d - 1)$, respectively.)

Simulating D_d

Given are x, y . Every round starts with a pair a_0, b_0 . If $f_{xy}(a_0) \neq (b_0)$ we have a certain error probability in simulating D_d in that round:

Corollary 2. *Given any x, y . In a round initialized with an incorrectly correlated pair a_0, b_0 , Alice and Bob simulate D_d , on inputs x, y , with an error probability of $2/d$.*

Proof. a_0, b_0 are incorrectly correlated if and only if $f_{xy}(a_0) \neq b_0$. The second step of the simulation round follows. The possible events are:

1. $s = 0$, where $P(s = 0) = 1/d$

According to (5,6) Alice and Bob locally compute $a_t = a_0, b_t = b_0$. Because $f_{xy}(a_0) \neq b_0 \Rightarrow f_{xy}(a) \neq b$ they obtain an incorrectly correlated pair a_t, b_t .

2. $s = 1$, where $P(s = 1) = 1 - 1/d$

According to (5,6) Alice and Bob locally compute $a_t = r_{a_0}(a)$, $b_t = r_{b_0}(b)$. Remember from (4) that the pairs $a = D_{d-1}^A(x')$, $b = D_{d-1}^B(y')$ always obey $g_{x'y'}(a) = b$, therefore $a_t = r_{a_0}(a)$ and $b_t = r_{b_0}(g_{x'y'}(a))$. $f_{xy}(a_0) \neq b_0$ implies $f_{xy}(a_0) = b^* \neq b_0$, $f_{xy}^{-1}(b_0) = a^* \neq a_0$. From the initialization of the round a_0, b_0 are fixed and therefore also a^*, b^* are fixed. Let a' denote any element in S_{d-1} different to a_0, a^* . We analyze the three possible assignments to a :

- $a = a^*$, where $P(a = a^*) = \frac{1}{d-1}$

$$\begin{aligned} a_t &= r_{a_0}(a^*) = a^*, \\ b_t &= r_{b_0}(g_{xa_0, yb_0}(a^*)) = r_{b_0}(r_{b_0}^{-1}(f_{xy}(a_0))) \\ &= f_{xy}(a_0) = b^*. \end{aligned}$$

- $a = a_0$, where $P(a = a_0) = \frac{1}{d-1}$

$$\begin{aligned} a_t &= r_{a_0}(a_0) = d - 1, \\ b_t &= r_{b_0}(g_{xa_0, yb_0}(a_0)) = r_{b_0}(r_{b_0}^{-1}(f_{xy}(r_{a_0}(a_0)))) \\ &= f_{xy}(r_{a_0}(a_0)) = f_{xy}(d - 1). \end{aligned}$$

- $a = a'$, where $P(a = a') = \frac{d-3}{d-1}$

$$\begin{aligned} a_t &= r_{a_0}(a') = a', \\ b_t &= r_{b_0}(g_{xa_0, yb_0}(a')) = r_{b_0}(r_{b_0}^{-1}(f_{xy}(r_{a_0}(a')))) \\ &= f_{xy}(r_{a_0}(a')) = f_{xy}(a'). \end{aligned}$$

We saw that if $s = 0$ the round will end in an incorrectly correlated pair, however, if $s = 1$ only the case $a = a^*$ yields an incorrect correlation, that is, $f_{xy}(a^*) \neq b^*$. The round ends with uniform probability in d different pairs a_t, b_t , such that $f_{xy}(a_t) = b_t$ in $d - 2$ cases and $f_{xy}(a_t) \neq b_t$ in 2 cases. That is, wrong output pairs — not correlated according to D_d given x, y — are obtained with a probability of $P(s = 0) + P(s = 1)P(a = a^*) = 2/d$. \square

The next corollary ensures that once we achieved a correct correlation in a round then all the following rounds will simulate D_d .

Corollary 3. *Given any x, y . In a round initialized with a correctly correlated pair a_0, b_0 , Alice and Bob simulate D_d on inputs x, y .*

Proof. a_0, b_0 are correctly correlated if and only if $f_{xy}(a_0) = b_0$. The second step of the simulation round follows. The possible events are:

1. $s = 0$, where $P(s = 0) = 1/d$

According to (5,6) Alice and Bob locally compute $a_t = a_0, b_t = b_0$. If $f_{xy}(a_0) = b_0$ then also $f_{xy}(a_t) = b_t$.

2. $s = 1$, where $P(s = 1) = 1 - 1/d$

According to (5,6) Alice and Bob locally compute $a_t = r_{a_0}(a)$, $b_t = r_{b_0}(b)$. Remember from (4) that the pairs $a = D_{d-1}^A(x')$, $b = D_{d-1}^B(y')$ always obey $g_{x'y'}(a) = b$, therefore $a_t = r_{a_0}(a)$; $b_t = r_{b_0}(g_{x'y'}(a))$. From the initialization of the round a_0, b_0 are fixed. Let a' denote any element in S_{d-1} different to a_0 . We analyze what happens locally in the two situations:

- $a = a_0$, where $P(a = a_0) = \frac{1}{d-1}$

$$\begin{aligned} a_t &= r_{a_0}(a_0) = d - 1, \\ b_t &= r_{b_0}(g_{xa_0, yb_0}(a_0)) = r_{b_0}(r_{b_0}^{-1}(f_{xy}(r_{a_0}(a_0)))) \\ &= f_{xy}(r_{a_0}(a_0)) = f_{xy}(d - 1). \end{aligned}$$

- $a = a'$, where $P(a = a') = \frac{d-2}{d-1}$

$$\begin{aligned} a_t &= r_{a_0}(a') = a', \\ b_t &= r_{b_0}(g_{xa_0, yb_0}(a')) = r_{b_0}(r_{b_0}^{-1}(f_{xy}(r_{a_0}(a')))) \\ &= f_{xy}(r_{a_0}(a')) = f_{xy}(a'). \end{aligned}$$

We see that the round ends with uniform probability in d different pairs a_t, b_t , such that always $f_{xy}(a_t) = b_t$, which is the exact behavior of D_d given x, y . \square

Induction Step

The shown construction of D_{d-1} from D_d and the proven corollaries of the simulation protocol are sufficient to complete this section by proving the following lemma:

Lemma 1. *For any alphabet size $d > 2$ any system $D_d \in \mathcal{D}$ can be simulated, within any error probability $\delta_d > 0$, by local operations on shared randomness and a finite number of approximations to a system $D_{d-1} \in \mathcal{D}$.*

Proof. Given is D_d . We construct D_{d-1} as previously defined. Let $\delta_{d-1} > 0$ denote its error probability, that is, the probability that outputs a, b on inputs x', y' do not satisfy $g_{x'y'}(a) = b$. According to Corollaries 2 and 3 the probability that our protocol simulates D_d after n rounds is therefore:

$$\Pr\{P_{D_d} = P_{\text{SIM}(n)}\} \geq (1 - \delta_{d-1})^n \left(\frac{d-2}{d}\right)^{n-1} \sum_{i=0}^{n-1} \left(\frac{2}{d}\right)^i. \quad (7)$$

Where $(1 - \delta_{d-1})^n$ denotes the probability that all the used systems D_{d-1} worked correctly. Obviously (7) is a geometric series with $q = \frac{2}{d}$. Therefore we can rewrite the term as

$$\Pr\{P_{D_d} = P_{\text{SIM}(n)}\} \geq (1 - \delta_{d-1})^n \left(1 - \left(\frac{2}{d}\right)^n\right).$$

We fix n as the, since d and δ_d are given, finite quantity

$$n = \varepsilon + \log_{2/d} \delta_d, \quad (8)$$

where $0 < \varepsilon \leq 1$ is used to round up to the next integer or maximally increase the directly obtained integer by one. We guarantee a maximal error probability δ_d by demanding

$$(1 - \delta_{d-1})^n \left(1 - \left(\frac{2}{d}\right)^n\right) = 1 - \delta_d.$$

By substituting n (8) we obtain

$$\begin{aligned} \delta_{d-1} &= 1 - \left(\frac{1 - \delta_d}{1 - (2/d)^n}\right)^{n^{-1}} \\ &= 1 - \left(\frac{1 - \delta_d}{1 - (2/d)^\varepsilon \delta_d}\right)^{(\varepsilon + \log_{2/d} \delta_d)^{-1}} > 0. \end{aligned}$$

Here $\varepsilon > 0$ and $d > 2$ ensure $\delta_{d-1} > 0$, that is, because $1 - \delta_d < 1 - (2/d)^\varepsilon \delta_d$ we can guarantee that there always exists a δ_{d-1} sufficiently small but still above zero, such that the error probability of our simulation is upper bounded by δ_d . \square

\mathcal{D} is universal

In this subsection we generalize our findings of the last subsection by proving that the set \mathcal{D} is actually universal in simulating any non-signaling correlation arbitrarily closely. More precisely, we will show the following:

Lemma 2. *For any system $P \in \mathcal{P}$ we find $D_d[P] \in \mathcal{D}$, such that Alice and Bob can approximate P by local operations on $D_d[P]$ arbitrarily well.*

Proof. Alice and Bob approximate P by simply locally relabeling the outputs of $D_d[P]$. According to Definition 2 it is sufficient to define $D_d[P]$ by the output alphabet $S_d = \{0, 1, \dots, d-1\}$, i.e., the cardinality d , and a set of permutations $\{f_{xy} : x \in X, y \in Y\}$ on S_d .

Constructing S_d

Given P we find d as follows: First we replace any non-signaling system P with *irrational* output probabilities $P(ab|xy) \in \mathbb{R} \setminus \mathbb{Q}$ with another non-signaling system which is entirely in the *rational number space* \mathbb{Q} and as close as desired to the original P . (Note that, as an implication of the following proof, any entirely rational P can be simulated perfectly by local operations on $D_d[P]$.) The parameter d is now chosen such that $1/d$ divides all output probabilities ($P(ab|xy)$ for all a, b, x, y) without remainder. So d shall be the *least common multiple* of the set of inverted output probabilities (divided by their denominators to have only integers).

Constructing D_d

Given P and S_d we find $\{f_{xy} : x \in X, y \in Y\}$ as follows: Let \mathcal{A}, \mathcal{B} denote P 's output sets of cardinalities d_a, d_b . Since $P \in \mathcal{P}$ the non-signaling conditions hold

$$\begin{aligned} \sum_{b \in \mathcal{B}} P(ab|xy) &= P(a|x) \text{ for all } a, x, y, \\ \sum_{a \in \mathcal{A}} P(ab|xy) &= P(b|y) \text{ for all } b, x, y. \end{aligned} \quad (9)$$

We fix arbitrary inputs x, y . First we partition the output set S_d into d_a and d_b pairwise disjoint parts, respectively

$$\begin{aligned} S_d &= A_{0x} \cup A_{1x} \cup \dots \cup A_{(d_a-1)x} = \bigcup_{a \in \mathcal{A}} A_{ax}, \\ S_d &= B_{0y} \cup B_{1y} \cup \dots \cup B_{(d_b-1)y} = \bigcup_{b \in \mathcal{B}} B_{by}. \end{aligned}$$

The parts shall have cardinalities

$$|A_{ax}| = dP(a|x) \text{ and } |B_{by}| = dP(b|y) \quad (10)$$

each. Next we partition S_d a second time into $d_a d_b$ disjoint parts

$$S_d = \bigcup_{a \in \mathcal{A}, b \in \mathcal{B}} A_{abxy}, \quad S_d = \bigcup_{a \in \mathcal{A}, b \in \mathcal{B}} B_{abxy}.$$

The parts shall have cardinalities

$$|A_{abxy}| = |B_{abxy}| = dP(ab|xy). \quad (11)$$

Note that if $P(ab|xy) = 0$ then $A_{abxy} = B_{abxy} = \emptyset$. Additionally the following conditions for all $a \in \mathcal{A}$ and $b \in \mathcal{B}$ shall hold:

$$A_{abxy} \subseteq A_{ax}, \quad B_{abxy} \subseteq B_{by}. \quad (12)$$

Observe that we can always choose such a second partitioning that fulfills the conditions (12), because

$$\sum_{b \in \mathcal{B}} |A_{abxy}| \stackrel{(11)}{=} \sum_{b \in \mathcal{B}} dP(ab|xy) \stackrel{(9)}{=} dP(a|x) \stackrel{(10)}{=} |A_{ax}|,$$

and therefore, for all a, x, y , we have $\sum_{b \in \mathcal{B}} |A_{abxy}| = |A_{ax}|$. The same argument obviously holds for Bobs second partitioning of S_d . The permutation f_{xy} is now defined by arbitrary bijections on these pairs of subsets, that is, given a, b ,

$$f_{xy} : A_{abxy} \rightarrow B_{abxy} \quad (13)$$

is defined and bijective. Since this holds for any pair a, b we have that f_{xy} is a bijection from S_d to itself.

Corollary 4. *Given x, y , f_{xy} is a permutation on S_d .*

Therefore the set of permutations $\{f_{xy} : x \in \mathcal{X}, y \in \mathcal{Y}\}$ defines a valid element of \mathcal{D} .

The Protocol

For the simulation, Alice and Bob have access to a system $D_d[P]$ and a description of the partitionings $\{A_{ax} : a \in \mathcal{A}, x \in \mathcal{X}\}$ and $\{B_{by} : b \in \mathcal{B}, y \in \mathcal{Y}\}$ respectively. For inputs x, y they locally obtain outputs a', b' from $D_d[P]$ indicated by functions $a' = D_d^A[P](x)$ and $b' = D_d^B[P](y)$ and select their final outputs a, b according to the rules

$$a \text{ if } a' \in A_{ax}, \quad b \text{ if } b' \in B_{by}. \quad (14)$$

The simulation protocol (SIM) is summarized in the following figure.

ALICE	SIM	BOB
$x \in \mathcal{X}$	inputs	$y \in \mathcal{Y}$
$\{A_{ax} : \forall a, x\}$ $a' = D_d^A[P](x)$	knowledge $D_d[P]$	$\{B_{by} : \forall b, y\}$ $b' = D_d^B[P](y)$
$a \text{ if } a' \in A_{ax}$	outputs	$b \text{ if } b' \in B_{by}$

FIG. 2: The simulation of P by local operations on one $D_d[P]$.

The Simulation

P will be simulated because for arbitrary x, y, a, b it follows, from the defined relabellings (14), that

$$P_{\text{SIM}}(ab|xy) \stackrel{(14)}{=} \sum_{a' \in A_{ax}, b' \in B_{by}} P_{D_d[P]}(a'b'|x, y).$$

Furthermore, the probability that $D_d[P]$ outputs a pair a', b' given a pair x, y is $1/d$ if and only if $f_{xy}(a') = b'$, therefore

$$\sum_{a' \in A_{ax}, b' \in B_{by}} P_{D_d[P]}(a'b'|x, y) = \sum_{\substack{a' \in A_{ax}, b' \in B_{by} \\ f_{xy}(a') = b'}} \frac{1}{d}.$$

We now concentrate on the summation index for a moment. By (13) we have the implication

$$f_{xy}(a') = b' \Rightarrow a' \in A_{a^*b^*xy}, \quad b' \in B_{a^*b^*xy}$$

for any a^*, b^* . The additional condition $a' \in A_{ax}, b' \in B_{by}$ in the summation index and (12) fixes $a^* = a, b^* = b$ to the originally given values. Therefore we can change the summation index as

$$\sum_{f_{xy}(a') = b', a' \in A_{ax}, b' \in B_{by}} \frac{1}{d} \stackrel{(13)}{=} \sum_{a' \in A_{abxy}, b' \in B_{abxy}} \frac{1}{d}.$$

The defined cardinalities $|A_{abxy}| = |B_{abxy}| = dP(ab|xy)$ in (11) lead to the wanted probability

$$\sum_{a' \in A_{abxy}, b' \in B_{abxy}} \frac{1}{d} = |A_{abxy}| \frac{1}{d} = |B_{abxy}| \frac{1}{d} = P(ab|xy).$$

Since a, b, x, y were arbitrary we have proved that the simulation holds for any outcome probability and therefore for P . \square

Let us summarize the situation: First we have proven the induction step from D_{d-1} to D_d , i.e., given a finite supply of imperfect systems D_{d-1} and shared randomness we constructed a non-interactive protocol between Alice and Bob that approximates D_d arbitrarily well. The induction is anchored by the fact that any D_2 , which is a distributed Boolean function, can be simulated by a finite number of non-local boxes [11]. Therefore, we have shown that a finite quantity of non-local boxes is sufficient for approximating any element of \mathcal{D} arbitrarily closely. In a second part we have provided a proof that \mathcal{D} is actually universal in approximating any non-signaling system. Concluding we have shown the claimed main result.

CONCLUDING REMARKS

With this note we have shown that any non-signaling bi-partite system can be approximated arbitrarily well by a finite number of non-local boxes. This complements an earlier result [10], which states that a *perfect* simulation is not always possible. Our result implies that the non-local box should be considered as a unit of bi-partite non-locality. Another implication is that every inner correlation, that is, any correlation which is not element of the convex hull of the non-signaling polytope, can be simulated perfectly by non-local boxes.

It is a challenging open problem whether the non-local box is also a universal unit for *multi*-partite non-signaling systems. If this fails to be true, a natural question would be whether there still exists a single system, or a finite number of systems, which is universal in the same sense as the NLB in the two-partite case.

This work was funded by the Swiss National Science Foundation (SNSF).

-
- [1] J. Bell, *Physics* **1**, 195 (1964).
 - [2] A. Einstein, B. Podolsky, and N. Rosen, *Physical Review* **47**, 777 (1935).
 - [3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press, 2000).
 - [4] S. Popescu and D. Rohrlich, *Foundations of Physics* **24**, 379 (1994).

- [5] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Physical Review Letters **23**, 880 (1969).
- [6] B. S. Tsirelson, Letters in Mathematical Physics **4**, 93 (1980).
- [7] N. S. Jones and L. Masanes, Physical Review **72**, 052312 (2005), quant-ph/0506182.
- [8] J. Barrett and S. Pironio, Physical Review Letters **95**, 140401 (2005), arXiv:quant-ph/0506180.
- [9] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Physical Review **71**, 022101 (2005), quant-ph/0404097.
- [10] F. Dupuis, N. Gisin, A. Hasidim, A. A. Méthot, and H. Pilpel, Journal of Mathematical Physics **48**, 082107 (2007), arXiv:quant-ph/0701142.
- [11] W. van Dam, ArXiv Quantum Physics e-prints (2005), arXiv:quant-ph/0501159.